Connecting via Winsock to STN

Welcome to STN International! Enter x:x

LOGINID:SSSPTA1635SXZ

PASSWORD:
TERMINAL (ENTER 1, 2, 3, OR ?):2

* * * * * * * * * *   Welcome to STN International * * * * * * * * * *

```
NEWS  1              Web Page URLs for STN Seminar Schedule - N. America
NEWS  2    Apr 08   "Ask CAS" for self-help around the clock
NEWS  3    Apr 09   BEILSTEIN: Reload and Implementation of a New Subject Area
NEWS  4    Apr 09   ZDB will be removed from STN
NEWS  5    Apr 19   US Patent Applications available in IFICDB, IFIPAT, and IFIUDB
NEWS  6    Apr 22   Records from IP.com available in CAPLUS, HCAPLUS, and ZCAPLUS
NEWS  7    Apr 22   BIOSIS Gene Names now available in TOXCENTER
NEWS  8    Apr 22   Federal Research in Progress (FEDRIP) now available
NEWS  9    Jun 03   New e-mail delivery for search results now available
NEWS 10    Jun 10   MEDLINE Reload
NEWS 11    Jun 10   PCTFULL has been reloaded
NEWS 12    Jul 02   FOREGE no longer contains STANDARDS file segment
NEWS 13    Jul 22   USAN to be reloaded July 28, 2002;
                    saved answer sets no longer valid
NEWS 14    Jul 29   Enhanced polymer searching in REGISTRY
NEWS 15    Jul 30   NETFIRST to be removed from STN
NEWS 16    Aug 08   CANCERLIT reload
NEWS 17    Aug 08   PHARMAMarketLetter(PHARMAML) - new on STN
NEWS 18    Aug 08   NTIS has been reloaded and enhanced
NEWS 19    Aug 19   Aquatic Toxicity Information Retrieval (AQUIRE)
                    now available on STN
NEWS 20    Aug 19   IFIPAT, IFICDB, and IFIUDB have been reloaded
NEWS 21    Aug 19   The MEDLINE file segment of TOXCENTER has been reloaded
NEWS 22    Aug 26   Sequence searching in REGISTRY enhanced
NEWS 23    Sep 03   JAPIO has been reloaded and enhanced
NEWS 24    Sep 16   Experimental properties added to the REGISTRY file
NEWS 25    Sep 16   CA Section Thesaurus available in CAPLUS and CA
NEWS 26    Oct 01   CASREACT Enriched with Reactions from 1907 to 1985
NEWS 27    Oct 21   EVENTLINE has been reloaded
NEWS 28    Oct 24   BEILSTEIN adds new search fields
NEWS 29    Oct 24   Nutraceuticals International (NUTRACEUT) now available on STN
NEWS 30    Oct 25   MEDLINE SDI run of October 8, 2002
NEWS 31    Nov 18   DKILIT has been renamed APOLLIT
NEWS 32    Nov 25   More calculated properties added to REGISTRY
NEWS 33    Dec 02   TIBKAT will be removed from STN
NEWS 34    Dec 04   CSA files on STN
NEWS 35    Dec 17   PCTFULL now covers WP/PCT Applications from 1978 to date
NEWS 36    Dec 17   TOXCENTER enhanced with additional content
NEWS 37    Dec 17   Adis Clinical Trials Insight now available on STN
NEWS 38    Dec 30   ISMEC no longer available
NEWS 39    Jan 21   NUTRACEUT offering one free connect hour in February 2003
NEWS 40    Jan 21   PHARMAML offering one free connect hour in February 2003
NEWS 41    Jan 29   Simultaneous left and right truncation added to COMPENDEX,
                    ENERGY, INSPEC
NEWS 42    Feb 13   CANCERLIT is no longer being updated
NEWS 43    Feb 24   METADEX enhancements
NEWS 44    Feb 24   PCTGEN now available on STN
NEWS 45    Feb 24   TEMA now available on STN
```

```
    NEWS 46   Feb 26   NTIS now allows simultaneous left and right truncation
.   NEWS·47   Feb 26   PCTFULL now contains images
    NEWS 48   Mar 04   SDI PACKAGE for monthly delivery of multifile SDI results
    NEWS 49   Mar 19   APOLLIT offering free connect time in April 2003
    NEWS 50   Mar 20   EVENTLINE will be removed from STN
    NEWS 51   Mar 24   PATDPAFULL now available on STN
    NEWS 52   Mar 24   Additional information for trade-named substances without
                       structures available in REGISTRY
    NEWS 53   Mar 24   Indexing from 1957 to 1966 added to records in CA/CAPLUS

    NEWS EXPRESS    April 4 CURRENT WINDOWS VERSION IS V6.01a, CURRENT
                    MACINTOSH VERSION IS V6.0b(ENG) AND V6.0Jb(JP),
                    AND CURRENT DISCOVER FILE IS DATED 01 APRIL 2003
    NEWS HOURS      STN Operating Hours Plus Help Desk Availability
    NEWS INTER      General Internet Information
    NEWS LOGIN      Welcome Banner and News Items
    NEWS PHONE      Direct Dial and Telecommunication Network Access to STN
    NEWS WWW        CAS World Wide Web Site (general information)

Enter NEWS followed by the item number or name to see news on that
specific topic.

  All use of STN is subject to the provisions of the STN Customer
  agreement.  Please note that this agreement limits use to scientific
  research.  Use for software development or design or implementation
  of commercial gateways or other similar uses is prohibited and may
  result in loss of user privileges and other penalties.

    * * * * * * * * * * * * * STN  Columbus  * * * * * * * * * * * * * * *

FILE 'HOME' ENTERED AT 09:27:37 ON 10 APR 2003

=> file .biotech
COST IN U.S. DOLLARS                              SINCE FILE      TOTAL
                                                    ENTRY        SESSION
FULL ESTIMATED COST                                 1.68          1.68

FILE 'MEDLINE' ENTERED AT 09:32:08 ON 10 APR 2003

FILE 'BIOSIS' ENTERED AT 09:32:08 ON 10 APR 2003
COPYRIGHT (C) 2003 BIOLOGICAL ABSTRACTS INC.(R)

FILE 'BIOTECHDS' ENTERED AT 09:32:08 ON 10 APR 2003
COPYRIGHT (C) 2003 THOMSON DERWENT AND INSTITUTE FOR SCIENTIFIC INFORMATION

FILE 'CAPLUS' ENTERED AT 09:32:08 ON 10 APR 2003
USE IS SUBJECT TO THE TERMS OF YOUR STN CUSTOMER AGREEMENT.
PLEASE SEE "HELP USAGETERMS" FOR DETAILS.
COPYRIGHT (C) 2003 AMERICAN CHEMICAL SOCIETY (ACS)

FILE 'EMBASE' ENTERED AT 09:32:08 ON 10 APR 2003
COPYRIGHT (C) 2003 Elsevier Science B.V.  All rights reserved.

=> s pars? (s) language
L1        364 PARS? (S) LANGUAGE

=> medic? or biolog?
MEDIC? IS NOT A RECOGNIZED COMMAND
The previous command name entered was not recognized by the system.
For a list of commands available to you in the current file, enter
"HELP COMMANDS" at an arrow prompt (=>).

=> s medic? or biolog?
   2 FILES SEARCHED...
```

L2   9166889 MEDIC? OR BIOLOG?

=> s l1 and l2
L3   117 L1 AND L2

=> d ti l3 1-20

L3   ANSWER 1 OF 117   MEDLINE
TI   The sublanguage of cross-coverage.

L3   ANSWER 2 OF 117   MEDLINE
TI   Finding UMLS Metathesaurus concepts in MEDLINE.

L3   ANSWER 3 OF 117   MEDLINE
TI   Variations in **Medical** Subject Headings (MeSH) mapping: from the
     natural language of patron terms to the controlled vocabulary of mapped
     lists.

L3   ANSWER 4 OF 117   MEDLINE
TI   Bantu **language** trees reflect the spread of farming across
     sub-Saharan Africa: a maximum-**parsimony** analysis.

L3   ANSWER 5 OF 117   MEDLINE
TI   SACS--self-maintaining database of antibody crystal structure information.

L3   ANSWER 6 OF 117   MEDLINE
TI   Automating SNOMED coding using **medical** language understanding: a
     feasibility study.

L3   ANSWER 7 OF 117   MEDLINE
TI   A knowledge model for the interpretation and visualization of NLP-parsed
     discharged summaries.

L3   ANSWER 8 OF 117   MEDLINE
TI   Meeting **medical** terminology needs--the Ontology-Enhanced
     Medical Concept Mapper.

L3   ANSWER 9 OF 117   MEDLINE
TI   XML for electronic clinical communications in Scotland.

L3   ANSWER 10 OF 117   MEDLINE
TI   Use of general-purpose negation detection to augment concept indexing of
     medical documents: a quantitative study using the UMLS.

L3   ANSWER 11 OF 117   MEDLINE
TI   Extracting clinical cases from XML-based electronic patient records for
     use in web-based **medical** case based reasoning systems.

L3   ANSWER 12 OF 117   MEDLINE
TI   Genetic and environmental risks for specific language impairment in
     children.

L3   ANSWER 13 OF 117   MEDLINE
TI   Semiotes: a semantics for sharing.

L3   ANSWER 14 OF 117   MEDLINE
TI   NLP techniques associated with the OpenGALEN ontology for semi-automatic
     textual extraction of **medical** knowledge: abstracting and mapping
     equivalent linguistic and logical constructs.

L3   ANSWER 15 OF 117   MEDLINE
TI   Limited parsing of notational text visit notes: ad-hoc vs. NLP approaches.

L3   ANSWER 16 OF 117   MEDLINE

TI   Semantic analysis of **medical** free texts.

L3   ANSWER 17 OF 117   MEDLINE
TI   The utilitarian core hypothesis: cases for testing the stability of
     languages in the wake of conquest.

L3   ANSWER 18 OF 117   MEDLINE
TI   EDGAR: extraction of drugs, genes and relations from the biomedical
     literature.

L3   ANSWER 19 OF 117   MEDLINE
TI   Language trees support the express-train sequence of Austronesian
     expansion.

L3   ANSWER 20 OF 117   MEDLINE
TI   A statistical natural language processor for **medical** reports.

=> dup rem 13
PROCESSING COMPLETED FOR L3
L4   90 DUP REM L3 (27 DUPLICATES REMOVED)

=> d ti 14 1-20

L4   ANSWER 1 OF 90 BIOSIS COPYRIGHT 2003 BIOLOGICAL ABSTRACTS INC.
TI   Anatomical correlates of dyslexia: Frontal and cerebellar findings.

L4   ANSWER 2 OF 90 BIOSIS COPYRIGHT 2003 BIOLOGICAL ABSTRACTS INC.
TI   Adaptive changes in early and late blind: A fMRI study of verb generation
     to heard nouns.

L4   ANSWER 3 OF 90 BIOSIS COPYRIGHT 2003 BIOLOGICAL ABSTRACTS INC.
TI   Do quiescent arachnoid cysts alter CNS functional organization? A fMRI and
     morphometric study.

L4   ANSWER 4 OF 90 BIOSIS COPYRIGHT 2003 BIOLOGICAL ABSTRACTS INC.
TI   Speech production: Wernicke, Broca and beyond.

L4   ANSWER 5 OF 90   MEDLINE   DUPLICATE 1
TI   Bantu **language** trees reflect the spread of farming across
     sub-Saharan Africa: a maximum-**parsimony** analysis.

L4   ANSWER 6 OF 90   MEDLINE
TI   The sublanguage of cross-coverage.

L4   ANSWER 7 OF 90   MEDLINE
TI   Finding UMLS Metathesaurus concepts in MEDLINE.

L4   ANSWER 8 OF 90 EMBASE COPYRIGHT 2003 ELSEVIER SCI. B.V.
TI   Novel metaphors appear anomalous at least momentarily: Evidence from N400.

L4   ANSWER 9 OF 90 BIOSIS COPYRIGHT 2003 BIOLOGICAL ABSTRACTS INC.
TI   Olfactory Receptor Database: A metadata-driven automated population from
     sources of gene and protein sequences.

L4   ANSWER 10 OF 90 BIOSIS COPYRIGHT 2003 BIOLOGICAL ABSTRACTS INC.
TI   ProML: The protein markup language for specification of protein sequences,
     structures and families.

L4   ANSWER 11 OF 90   MEDLINE   DUPLICATE 2
TI   SACS--self-maintaining database of antibody crystal structure information.

L4   ANSWER 12 OF 90   MEDLINE
TI   Variations in **Medical** Subject Headings (MeSH) mapping: from the

natural language of patron terms to the controlled vocabulary of mapped lists.

L4   ANSWER 13 OF 90   MEDLINE
TI   Extracting clinical cases from XML-based electronic patient records for use in web-based **medical** case based reasoning systems.

L4   ANSWER 14 OF 90   MEDLINE   DUPLICATE 3
TI   Use of general-purpose negation detection to augment concept indexing of **medical** documents: a quantitative study using the UMLS.

L4   ANSWER 15 OF 90   MEDLINE
TI   Automating SNOMED coding using **medical** language understanding: a feasibility study.

L4   ANSWER 16 OF 90   MEDLINE   DUPLICATE 4
TI   XML for electronic clinical communications in Scotland.

L4   ANSWER 17 OF 90   MEDLINE   DUPLICATE 5
TI   Genetic and environmental risks for specific language impairment in children.

L4   ANSWER 18 OF 90   MEDLINE
TI   A knowledge model for the interpretation and visualization of NLP-parsed discharged summaries.

L4   ANSWER 19 OF 90   MEDLINE
TI   Meeting **medical** terminology needs--the Ontology-Enhanced **Medical** Concept Mapper.

L4   ANSWER 20 OF 90 BIOSIS COPYRIGHT 2003 BIOLOGICAL ABSTRACTS INC.
TI   A psycholinguistically and neurolinguistically plausible computational model of natural-language processing by the human brain.

=> d ti 14 21-40

L4   ANSWER 21 OF 90 BIOSIS COPYRIGHT 2003 BIOLOGICAL ABSTRACTS INC.
TI   Anomalous anatomy of speech-language areas in adults with persistent developmental stuttering.

L4   ANSWER 22 OF 90 BIOSIS COPYRIGHT 2003 BIOLOGICAL ABSTRACTS INC.
TI   Past tense morphology in specifically language impaired and normally developing children.

L4   ANSWER 23 OF 90 BIOSIS COPYRIGHT 2003 BIOLOGICAL ABSTRACTS INC.
TI   The voice of historical biogeography.

L4   ANSWER 24 OF 90 BIOTECHDS COPYRIGHT 2003 THOMSON DERWENT AND ISI
TI   Identifying novel nucleic acid molecules encoding proteins of interest, and natural language processing and extraction of relational information associated with genes and proteins found in journal articles; method is useful for identifying novel nucleic acid molecule

L4   ANSWER 25 OF 90 BIOSIS COPYRIGHT 2003 BIOLOGICAL ABSTRACTS INC.
TI   Interhemispheric transfer of language in patients with left frontal cerebral arteriovenous malformation.

L4   ANSWER 26 OF 90   MEDLINE   DUPLICATE 6
TI   Semiotes: a semantics for sharing.

L4   ANSWER 27 OF 90   MEDLINE   DUPLICATE 7
TI   Language trees support the express-train sequence of Austronesian expansion.

```
L4    ANSWER 28 OF 90  BIOSIS  COPYRIGHT 2003 BIOLOGICAL ABSTRACTS INC.
TI    Segregating semantic and syntactic aspects of processing in the human
      brain: An fMRI investigation of different word types.

L4    ANSWER 29 OF 90      MEDLINE
TI    The utilitarian core hypothesis: cases for testing the stability of
      languages in the wake of conquest.

L4    ANSWER 30 OF 90      MEDLINE
TI    EDGAR: extraction of drugs, genes and relations from the biomedical
      literature.

L4    ANSWER 31 OF 90      MEDLINE
TI    Semantic analysis of medical free texts.

L4    ANSWER 32 OF 90  BIOSIS  COPYRIGHT 2003 BIOLOGICAL ABSTRACTS INC.
TI    Induction of a marsupial density model using genetic programming and
      spatial relationships.

L4    ANSWER 33 OF 90      MEDLINE
TI    NLP techniques associated with the OpenGALEN ontology for semi-automatic
      textual extraction of medical knowledge: abstracting and mapping
      equivalent linguistic and logical constructs.

L4    ANSWER 34 OF 90      MEDLINE
TI    Limited parsing of notational text visit notes: ad-hoc vs. NLP approaches.

L4    ANSWER 35 OF 90  BIOSIS  COPYRIGHT 2003 BIOLOGICAL ABSTRACTS INC.
TI    Testing the generalized slowing hypothesis in specific language
      impairment.

L4    ANSWER 36 OF 90  BIOSIS  COPYRIGHT 2003 BIOLOGICAL ABSTRACTS INC.
TI    Language related brain potentials in patients with cortical and
      subcortical left hemisphere lesions.

L4    ANSWER 37 OF 90      MEDLINE
TI    A statistical natural language processor for medical reports.

L4    ANSWER 38 OF 90      MEDLINE
TI    Extracting noun phrases for all of MEDLINE.

L4    ANSWER 39 OF 90  EMBASE  COPYRIGHT 2003 ELSEVIER SCI. B.V.
TI    MERIT-9: A patient information exchange guideline.

L4    ANSWER 40 OF 90      MEDLINE                          DUPLICATE 8
TI    Representing information in patient reports using natural language
      processing and the extensible markup language.


=> d ti 14 41-90

L4    ANSWER 41 OF 90      MEDLINE
TI    Patient information exchange guideline MERIT-9 using medical
      markup language MML.

L4    ANSWER 42 OF 90  BIOSIS  COPYRIGHT 2003 BIOLOGICAL ABSTRACTS INC.
TI    Rule-based management for simulation in agricultural decision support
      systems.

L4    ANSWER 43 OF 90      MEDLINE                          DUPLICATE 9
TI    Dependency parsing for medical language and
      concept representation.
```

L4   ANSWER 44 OF 90   MEDLINE   DUPLICATE 10
TI   MERIT-9: a patient information exchange guideline using MML, HL7 and
      DICOM.

L4   ANSWER 45 OF 90   BIOSIS   COPYRIGHT 2003 BIOLOGICAL ABSTRACTS INC.
TI   Consciousness in neural networks.

L4   ANSWER 46 OF 90   MEDLINE
TI   A natural **language parsing** system for encoding
      admitting diagnoses.

L4   ANSWER 47 OF 90   BIOSIS   COPYRIGHT 2003 BIOLOGICAL ABSTRACTS INC.
TI   Cannabinoid receptors in the human brain: A detailed anatomical and
      quantitative autoradiographic study in the fetal, neonatal and adult human
      brain.

L4   ANSWER 48 OF 90   BIOSIS   COPYRIGHT 2003 BIOLOGICAL ABSTRACTS INC.
TI   **Pars** triangularis asymmetry and **language** dominance.

L4   ANSWER 49 OF 90   BIOSIS   COPYRIGHT 2003 BIOLOGICAL ABSTRACTS INC.
TI   Corticobasal degeneration with primary progressive aphasia and accentuated
      cortical lesion in superior temporal gyrus: Case report and review.

L4   ANSWER 50 OF 90   MEDLINE
TI   Recognizing noun phrases in **medical** discharge summaries: an
      evaluation of two natural **language parsers**.

L4   ANSWER 51 OF 90   MEDLINE
TI   Toward reusable software components at the point of care.

L4   ANSWER 52 OF 90   BIOSIS   COPYRIGHT 2003 BIOLOGICAL ABSTRACTS INC.
TI   New trends in natural language processing: Statistical natural language
      processing.

L4   ANSWER 53 OF 90   BIOSIS   COPYRIGHT 2003 BIOLOGICAL ABSTRACTS INC.
TI   Functional MRI measurement of language lateralization in Wada-tested
      patients.

L4   ANSWER 54 OF 90   BIOSIS   COPYRIGHT 2003 BIOLOGICAL ABSTRACTS INC.
TI   Automated **parsing** of natural **language** text data from
      death certificates and other sources.

L4   ANSWER 55 OF 90   BIOSIS   COPYRIGHT 2003 BIOLOGICAL ABSTRACTS INC.
TI   Use of Fluconazole in the Treatment of Candidal Endophthalmitis.

L4   ANSWER 56 OF 90   BIOSIS   COPYRIGHT 2003 BIOLOGICAL ABSTRACTS INC.
TI   Distribution of the four founding lineage haplotypes in native Americans
      suggests a single wave of migration for the New World.

L4   ANSWER 57 OF 90   BIOSIS   COPYRIGHT 2003 BIOLOGICAL ABSTRACTS INC.
TI   Time estimation deficits in developmental dyslexia: Evidence of cerebellar
      involvement.

L4   ANSWER 58 OF 90   BIOSIS   COPYRIGHT 2003 BIOLOGICAL ABSTRACTS INC.
TI   Disturbances of learning processes in the basal ganglia in the
      pathogenesis of Parkinson's disease: A novel theory.

L4   ANSWER 59 OF 90   MEDLINE
TI   Associating semantic grammars with the SNOMED: processing **medical**
      language and representing clinical facts into a language-independent
      frame.

L4   ANSWER 60 OF 90   BIOSIS   COPYRIGHT 2003 BIOLOGICAL ABSTRACTS INC.
TI   A Darwinian approach to the origins of psychosis.

```
L4    ANSWER 61 OF 90    MEDLINE                        DUPLICATE 11
TI    Macromolecular query language (MMQL): prototype data model and
      implementation.

L4    ANSWER 62 OF 90    MEDLINE
TI    A natural language understanding system combining syntactic and semantic
      techniques.

L4    ANSWER 63 OF 90    MEDLINE
TI    A general natural-language text processor for clinical radiology.

L4    ANSWER 64 OF 90  BIOSIS  COPYRIGHT 2003 BIOLOGICAL ABSTRACTS INC.
TI    The anatomy of an ecological controversy: Honey-bee searching behavior.

L4    ANSWER 65 OF 90    MEDLINE                        DUPLICATE 12
TI    Users conceptual views on medical information databases.

L4    ANSWER 66 OF 90    MEDLINE
TI    Design and application of a C++ macromolecular class library.

L4    ANSWER 67 OF 90    MEDLINE
TI    A conceptual model for information retrieval with UMLS.

L4    ANSWER 68 OF 90    MEDLINE
TI    Generating MEDLINE search strategies using a librarian knowledge-based
      system.

L4    ANSWER 69 OF 90    MEDLINE
TI    Interpreting natural language queries using the UMLS.

L4    ANSWER 70 OF 90    MEDLINE
TI    Computer auditing of surgical operative reports written in English.

L4    ANSWER 71 OF 90    MEDLINE
TI    UMLS knowledge for biomedical language processing.

L4    ANSWER 72 OF 90    MEDLINE                        DUPLICATE 13
TI    Semantic analysis of medical records.

L4    ANSWER 73 OF 90    MEDLINE
TI    A history-taking system that uses continuous speech recognition.

L4    ANSWER 74 OF 90    MEDLINE
TI    The role of automated speech recognition in endoscopic data collection.

L4    ANSWER 75 OF 90    MEDLINE                        DUPLICATE 14
TI    Evaluation of a Meta-1-based automatic indexing method for medical
      documents.

L4    ANSWER 76 OF 90    MEDLINE                        DUPLICATE 15
TI    Natural language processing and semantical representation of
      medical texts.

L4    ANSWER 77 OF 90    MEDLINE
TI    An automatic indexing method for medical documents.

L4    ANSWER 78 OF 90    MEDLINE                        DUPLICATE 16
TI    A Medical Text Analysis System for German--syntax analysis.

L4    ANSWER 79 OF 90    MEDLINE
TI    Extending a natural language parser with UMLS
      knowledge.
```

L4   ANSWER 80 OF 90   EMBASE   COPYRIGHT 2003 ELSEVIER SCI. B.V.
TI   Computer analysis of sublanguage information structures.

L4   ANSWER 81 OF 90   MEDLINE   DUPLICATE 17
TI   A prototype system for perinatal knowledge engineering using an artificial
      intelligence tool.

L4   ANSWER 82 OF 90   MEDLINE
TI   Locative inferences in **medical** texts.

L4   ANSWER 83 OF 90   MEDLINE
TI   **Biological** processing.

L4   ANSWER 84 OF 90   EMBASE   COPYRIGHT 2003 ELSEVIER SCI. B.V.
TI   **Biological** processing.

L4   ANSWER 85 OF 90   MEDLINE   DUPLICATE 18
TI   Automatic encoding of clinical narrative.

L4   ANSWER 86 OF 90   EMBASE   COPYRIGHT 2003 ELSEVIER SCI. B.V.
TI   Profile of a dictionary compiled from scanning over one million words of
      surgical pathology narrative text.

L4   ANSWER 87 OF 90   EMBASE   COPYRIGHT 2003 ELSEVIER SCI. B.V.
TI   A redefinition of the syndrome of Broca's aphasia: implications for a
      neuropsychological model of language.

L4   ANSWER 88 OF 90   EMBASE   COPYRIGHT 2003 ELSEVIER SCI. B.V.
TI   Data security in information systems by language analysis.

L4   ANSWER 89 OF 90   BIOSIS   COPYRIGHT 2003 BIOLOGICAL ABSTRACTS INC.
TI   CLINICAL PSYCHO PHYSICS APPLICATIONS OF RATIO SCALING AND SIGNAL DETECTION
      METHODS TO RESEARCH ON PAIN FEAR DRUGS AND **MEDICAL** DECISION
      MAKING.

L4   ANSWER 90 OF 90   EMBASE   COPYRIGHT 2003 ELSEVIER SCI. B.V.
TI   Language processor generation with BNF inputs: Mehods and implementation.

=> d ibib abs 14 10, 12-15, 18, 24, 30, 31, 33, 34, 37, 38, 40, 43, 46, 50, 52, 62,
63, 67-69, 71, 72, 76-80, 83, 84, 86

L4   ANSWER 10 OF 90   BIOSIS   COPYRIGHT 2003 BIOLOGICAL ABSTRACTS INC.
ACCESSION NUMBER: 2003:124437 BIOSIS
DOCUMENT NUMBER: PREV200300124437
TITLE: ProML: The protein markup language for specification of
                     protein sequences, structures and families.
AUTHOR(S):   Hanisch, Daniel (1); Zimmer, Ralf; Lengauer, Thomas
CORPORATE SOURCE: (1) Fraunhofer Institute for Algorithms and Scientific
                     Computing (SCAI), Schloss Birlinghoven, D-53754, Sankt
                     Augustin, Germany Germany
SOURCE: In Silico Biology, (2002) Vol. 2, No. 3, pp. 313-324.
                     print.
                     ISSN: 1386-6338.
DOCUMENT TYPE: Article
LANGUAGE: English
AB   We propose a specification **language** ProML for protein sequences,
      structures, and families based on the open XML standard. The
      language allows for portable, system-independent, machine-
      parsable and human-readable representation of essential features
      of proteins. The **language** is of immediate use for several
      bioinformatics applications: we discuss clustering of proteins into
      families and the representation of the specific shared features of the
      respective clusters. Moreover, we use ProML for specification of data used

in fold recognition bench-marks exploiting experimentally derived distance constraints.

ACCESSION NUMBER: 2002259392   MEDLINE
DOCUMENT NUMBER: 21993818   PubMed ID: 11999175
TITLE: Variations in **Medical** Subject Headings (MeSH)
                     mapping: from the natural language of patron terms to the
                     controlled vocabulary of mapped lists.
COMMENT: Comment in: J Med Libr Assoc. 2002 Oct;90(4):475
AUTHOR: Gault Lora V; Shultz Mary; Davies Kathy J
CORPORATE SOURCE: The Library, Purdue University Calumet, Hammond, Indiana
                     46323-2590, USA.. gault@calumet.purdue.edu
SOURCE: J Med Libr Assoc, (2002 Apr) 90 (2) 173-80.
                     Journal code: 101132728. ISSN: 1536-5050.
PUB. COUNTRY: United States
DOCUMENT TYPE: Journal; Article; (JOURNAL ARTICLE)
LANGUAGE: English
FILE SEGMENT: Priority Journals
ENTRY MONTH: 200210
ENTRY DATE: Entered STN: 20020510
                     Last Updated on STN: 20030215
                     Entered Medline: 20021022

AB   OBJECTIVES: This study compared the mapping of natural **language**
      patron terms to the **Medical** Subject Headings (MeSH) across six
      MeSH interfaces for the MEDLINE database. METHODS: Test data were obtained
      from search requests submitted by patrons to the Library of the Health
      Sciences, University of Illinois at Chicago, over a nine-month period.
      Search request statements were **parsed** into separate terms or
      phrases. Using print sources from the National Library of **Medicine**
      , Each **parsed** patron term was assigned corresponding MeSH terms.
      Each patron term was entered into each of the selected interfaces to
      determine how effectively they mapped to MeSH. Data were collected for
      mapping success, accessibility of MeSH term within mapped list, and total
      number of MeSH choices within each list. RESULTS: The selected MEDLINE
      interfaces do not map the same patron term in the same way, nor do they
      consistently lead to what is considered the appropriate MeSH term.
      CONCLUSIONS: If searchers utilize the MEDLINE database to its fullest
      potential by mapping to MeSH, the results of the mapping will vary between
      interfaces. This variance may ultimately impact the search results. These
      differences should be considered when choosing a MEDLINE interface and
      when instructing end users.

ACCESSION NUMBER: 2001556397   MEDLINE
DOCUMENT NUMBER: 21490603   PubMed ID: 11604816
TITLE: Extracting clinical cases from XML-based electronic patient
                     records for use in web-based **medical** case based
                     reasoning systems.
AUTHOR: Manickam S; Abidi S S
CORPORATE SOURCE: School of Computer Sciences,Universiti Sains Malaysia,11800
                     Penang, Malaysia.. selva@cs.usm.my
SOURCE: MEDINFO, (2001) 10 (Pt 1) 643-7.
                     Journal code: 7600347. ISSN: 1569-6332.
PUB. COUNTRY: Netherlands
DOCUMENT TYPE: Journal; Article; (JOURNAL ARTICLE)
LANGUAGE: English
FILE SEGMENT: Priority Journals
ENTRY MONTH: 200201
ENTRY DATE: Entered STN: 20011018
                     Last Updated on STN: 20020125
                     Entered Medline: 20020108

AB   Development and usage of Case Based Reasoning (CBR) driven **medical**
      diagnostic system requires a large volume of clinical cases that depict

the problem-solving methodology of **medical** experts. Successful
usage of CBR based systems in healthcare is constrained by the need for a
continuous supply of current and correct clinical cases (in an electronic
medium) from **medical** experts. To address this constraint we
present a strategy to pro-actively transform generic Electronic Patient
Records (EPR) to Operable CBR-oriented Cases (OCC) that are compliant to
specialised CBR-based **medical** systems. EPR-OCC transformation
methodology is based on XML **parse**-trees, Unified **Medical**
Language Source (UMLS) meta-thesauri and **medical**
knowledge ontologies. The featured work involves the implementation of a
Java-based computer system for the automatic transformation of XML-based
EPR-originating from heterogeneous EPR repositories accessible over the
Internet/WWW-to specialised OCC that can then be seamlessly incorporated
within Intelligent CBR-based **Medical** Diagnostic Systems.

L4   ANSWER 14 OF 90   MEDLINE   DUPLICATE 3
ACCESSION NUMBER: 2001639139   MEDLINE
DOCUMENT NUMBER: 21547134   PubMed ID: 11687566
TITLE: Use of general-purpose negation detection to augment
                     concept indexing of **medical** documents: a
                     quantitative study using the UMLS.
AUTHOR: Mutalik P G; Deshpande A; Nadkarni P M
CORPORATE SOURCE: Department of Diagnostic Raiology, Yale University School
                     of Medicine, New Haven, Connecticut 06510, USA..
                     Pradeep.Mutalik@yale.edu
CONTRACT NUMBER: R01 LM06843-01 (NLM)
SOURCE: JOURNAL OF THE AMERICAN MEDICAL INFORMATICS ASSOCIATION,
                     (2001 Nov-Dec) 8 (6) 598-609.
                     Journal code: 9430800. ISSN: 1067-5027.
PUB. COUNTRY: United States
DOCUMENT TYPE: (EVALUATION STUDIES)
                     Journal; Article; (JOURNAL ARTICLE)
LANGUAGE: English
FILE SEGMENT: Priority Journals
ENTRY MONTH: 200112
ENTRY DATE: Entered STN: 20011107
                     Last Updated on STN: 20020123
                     Entered Medline: 20011211

AB   OBJECTIVES: To test the hypothesis that most instances of negated concepts
     in dictated **medical** documents can be detected by a strategy that
     relies on tools developed for the **parsing** of formal (computer)
     languages-specifically, a lexical scanner ("lexer") that uses regular
     expressions to generate a finite state machine, and a **parser**
     that relies on a restricted subset of context-free grammars, known as
     LALR(1) grammars. METHODS: A diverse training set of 40 **medical**
     documents from a variety of specialties was manually inspected and used to
     develop a program (Negfinder) that contained rules to recognize a large
     set of negated patterns occurring in the text. Negfinder's lexer and
     parser were developed using tools normally used to generate
     programming **language** compilers. The input to Negfinder consisted
     of **medical** narrative that was preprocessed to recognize UMLS
     concepts: the text of a recognized concept had been replaced with a coded
     representation that included its UMLS concept ID. The program generated an
     index with one entry per instance of a concept in the document, where the
     presence or absence of negation of that concept was recorded. This
     information was used to mark up the text of each document by color-coding
     it to make it easier to inspect. The **parser** was then evaluated
     in two ways: 1) a test set of 60 documents (30 discharge summaries, 30
     surgical notes) marked-up by Negfinder was inspected visually to quantify
     false-positive and false-negative results; and 2) a different test set of
     10 documents was independently examined for negatives by a human observer
     and by Negfinder, and the results were compared. RESULTS: In the first
     evaluation using marked-up documents, 8,358 instances of UMLS concepts
     were detected in the 60 documents, of which 544 were negations detected by

the program and verified by human observation (true-positive results, or TPs). Thirteen instances were wrongly flagged as negated (false-positive results, or FPs), and the program missed 27 instances of negation (false-negative results, or FNs), yielding a sensitivity of 95.3 percent and a specificity of 97.7 percent. In the second evaluation using independent negation detection, 1,869 concepts were detected in 10 documents, with 135 TPs, 12 FPs, and 6 FNs, yielding a sensitivity of 95.7 percent and a specificity of 91.8 percent. One of the words "no," "denies/denied," "not," or "without" was present in 92.5 percent of all negations. CONCLUSIONS: Negation of most concepts in **medical** narrative can be reliably detected by a simple strategy. The reliability of detection depends on several factors, the most important being the accuracy of concept matching.

L4   ANSWER 15 OF 90   MEDLINE
ACCESSION NUMBER: 2002094475   MEDLINE
DOCUMENT NUMBER: 21684218   PubMed ID: 11825222
TITLE: Automating SNOMED coding using **medical** language understanding: a feasibility study.
AUTHOR: Lussier Y A; Shagina L; Friedman C
CORPORATE SOURCE: Department of medical Informatics, Columbia University, New York, NY, USA.
CONTRACT NUMBER: LM06274 (NLM)
SOURCE: PROCEEDINGS / AMIA SYMPOSIUM, (2001) 418-22.
    Journal code: 100883449. ISSN: 1531-605X.
PUB. COUNTRY: United States
DOCUMENT TYPE: (EVALUATION STUDIES)
    Journal; Article; (JOURNAL ARTICLE)
LANGUAGE: English
FILE SEGMENT: Priority Journals
ENTRY MONTH: 200205
ENTRY DATE: Entered STN: 20020205
    Last Updated on STN: 20020528
    Entered Medline: 20020524
AB   This paper evaluates qualitatively the use of the MedLEE natural **language** processing system to code **medical** narratives directly into the SNOMED nomenclature, while retaining the MedLEE information model data structure. A gold standard is produced from narrative text manually coded in SNOMED. An automated **parsing** and SNOMED-coding of the narrative text is then automatically generated by MedLEE. By comparing MedLEE s output to that of the Gold Standard, the capacities of SNOMED and MedLEE to represent the clinical information are subsequently evaluated leading to qualitative observations on their respective strengths and constraints. In this study, MedLEE did code to SNOMED and captures the codes in a sub-structure amenable to interoperability with the description logic of SNOMED RT, showing an approach that augments and formalizes SNOMED s compositional representation methods to accurately capture information from clinical narratives.

L4   ANSWER 18 OF 90   MEDLINE
ACCESSION NUMBER: 2002094460   MEDLINE
DOCUMENT NUMBER: 21684203   PubMed ID: 11825207
TITLE: A knowledge model for the interpretation and visualization of NLP-parsed discharged summaries.
AUTHOR: Krauthammer M; Hripcsak G
CORPORATE SOURCE: Medical Informatics, Columbia University, New York, NY, USA.
CONTRACT NUMBER: R01-LM06274 (NLM)
    R01-LM06910 (NLM)
SOURCE: PROCEEDINGS / AMIA SYMPOSIUM, (2001) 339-43.
    Journal code: 100883449. ISSN: 1531-605X.
PUB. COUNTRY: United States
DOCUMENT TYPE: (EVALUATION STUDIES)

AB   At our institution, a Natural **Language** Processing (NLP) tool
     called MedLEE is used on a daily basis to **parse medical**
     texts including complete discharge summaries. MedLEE transforms written
     text into a generic structured format, which preserves the richness of the
     underlying natural **language** expressions by the use of concept
     modifiers (like change, certainty, degree and status). As a tradeoff,
     extraction of application-specific **medical** information is
     difficult without a clear understanding of how these modifiers combine. We
     report on a knowledge model for MedLEE modifiers that is helpful for a
     high level interpretation of NLP data and is used for the generation of
     two distinct views on NLP-**parsed** discharge summaries: A
     physician view offering a condensed overview of the severity of patient
     problems and a data mining view featuring binary problem states useful for
     machine learning.

L4   ANSWER 24 OF 90   BIOTECHDS COPYRIGHT 2003 THOMSON DERWENT AND ISI
ACCESSION NUMBER: 2001-01426   BIOTECHDS
TITLE: Identifying novel nucleic acid molecules encoding proteins of
                    interest, and natural language processing and extraction of
                    relational information associated with genes and proteins
                    found in journal articles;
                        method is useful for identifying novel nucleic acid
                        molecule
AUTHOR: Rzhetsky A; Kalachikov S; Krauthammer M O; Friedman C; Kra P
PATENT ASSIGNEE: Univ.New-York-Columbia
LOCATION: New York, NY, USA.
PATENT INFO: WO 2000063687 26 Oct 2000
APPLICATION INFO: WO 2000-US10302 14 Apr 2000
PRIORITY INFO: US 1999-327983 8 Jun 1999; US 1999-129469 15 Apr 1999
DOCUMENT TYPE: Patent
LANGUAGE: English
OTHER SOURCE: WPI: 2000-665269 [64]
AN   2001-01426   BIOTECHDS
AB   A new method for identifying novel nucleic acid molecules encoding a
       protein of interest, using regulatory networks is claimed. Also claimed
       are: identifying the effect of a gene knockout on a regulatory pathway;
       identifying a novel nucleic acid molecule encoding a protein of interest;
       identifying a novel gene; extracting information on interactions between
       biological entities from natural-**language** text data; a
       computer system for extracting information on **biological**
       entities from natural-**language** text data. The method further
       involves using each identified expression sequence tag to search sequence
       databases for overlapping sequences, to assemble longer overlapping
       stretches of DNA. The method further involves preprocessing, the data
       prior to **parsing**. The method is useful for identifying novel
       genes and for natural **language** processing and extraction of
       relational information associated with genes and proteins that are found
       in genomics journal articles. (374pp)

L4   ANSWER 30 OF 90   MEDLINE
ACCESSION NUMBER: 2000410553   MEDLINE
DOCUMENT NUMBER: 20360656   PubMed ID: 10902199
TITLE: EDGAR: extraction of drugs, genes and relations from the
                     biomedical literature.
AUTHOR: Rindflesch T C; Tanabe L; Weinstein J N; Hunter L
CORPORATE SOURCE: Lister Hill Center, National Library of Medicine, Bethesda,
                     MD 20894, USA.. tcr@lhc.nlm.nih.gov

AB   EDGAR (Extraction of Drugs, Genes and Relations) is a natural
      language processing system that extracts information about drugs
      and genes relevant to cancer from the biomedical literature. This
      automatically extracted information has remarkable potential to facilitate
      computational analysis in the molecular **biology** of cancer, and
      the technology is straightforwardly generalizable to many areas of
      biomedicine. This paper reports on the mechanisms for automatically
      generating such assertions and on a simple application, conceptual
      clustering of documents. The system uses a stochastic part of speech
      tagger, generates an underspecified syntactic **parse** and then
      uses semantic and pragmatic information to construct its assertions. The
      system builds on two important existing resources: the MEDLINE database of
      biomedical citations and abstracts and the Unified **Medical**
      Language System, which provides syntactic and semantic information
      about the terms found in biomedical abstracts.

AB   The semantic interpretation of natural **language** utterances is
      usually based on a large number of transformation rules which map
      syntactic structures (**parse** trees) onto some kind of meaning
      representation. However, those interpretation rules exhibit an
      insufficient degree of abstraction so that the scalability and portability
      of such natural **language** processing systems is hard to maintain.
      In this paper, we introduce an approach that is able to cope with a wide
      variety of semantic interpretation patterns in **medical** free
      texts by applying a small inventory of abstract semantic interpretation
      schemata. These schemata address generalized graph configurations within
      syntactic dependency **parse** trees, which abstract away from
      specific syntactic constructions.

AUTHOR: do Amaral M B; Roberts A; Rector A L
CORPORATE SOURCE: Sao Paulo University, Medical School Hospital, Brazil.

AB   This research project presents methodological and theoretical issues
     related to the inter-relationship between linguistic and conceptual
     semantics, analysing the results obtained by the application of a NLP
     parser to a set of radiology reports. Our objective is to define a
     technique for associating linguistic methods with domain specific
     ontologies for semi-automatic extraction of intermediate representation
     (IR) information formats and **medical** ontological knowledge from
     clinical texts. We have applied the Edinburgh LTG natural **language**
     parser to 2810 clinical narratives describing radiology
     procedures. In a second step, we have used **medical** expertise and
     ontology formalism for identification of semantic structures and
     abstraction of IR schemas related to the processed texts. These IR schemas
     are an association of linguistic and conceptual knowledge, based on their
     semantic contents. This methodology aims to contribute to the elaboration
     of models relating linguistic and logical constructs based on empirical
     data analysis. Advance in this field might lead to the development of
     computational techniques for automatic enrichment of **medical**
     ontologies from real clinical environments, using descriptive knowledge
     implicit in large text corpora sources.

TITLE: Limited parsing of notational text visit notes: ad-hoc vs.
                    NLP approaches.
AUTHOR: Barrows Jr R C; Busuioc M; Friedman C
CORPORATE SOURCE: Department of Medical Informatics, Columbia University, New
                    York, NY, USA.

AB   This paper describes the extraction of structured data relevant to
     glaucoma diagnosis and progression from visit notes typed as "notational
     text" by ophthalmologists during patient encounters. We compared two text
     processing systems: a limited pattern matching system called GDP (Glaucoma
     Dedicated **Parser**) and MedLEE, a proven natural **language**
     processing system which is in routine use encoding findings from chest
     radiograph and mammogram reports at the New York-Presbyterian hospital's
     Columbia-Presbyterian Center. We also evaluated the use of GDP as a
     preprocessor program to transform notational text into constructions
     recognizable by MedLEE. These systems have been evaluated according to
     their recall and precision in the particular task of processing a corpus
     of "notational text" documents to extract information related to glaucoma
     disease.

ACCESSION NUMBER: 2000033015   MEDLINE
DOCUMENT NUMBER: 20033015   PubMed ID: 10566505
TITLE: A statistical natural language processor for
                       medical reports.
AUTHOR: Taira R K; Soderland S G
CORPORATE SOURCE: Department of Radiology, Children's Hospital, Seattle, WA
                       98105, USA.
SOURCE: PROCEEDINGS / AMIA SYMPOSIUM, (1999) 970-4.
                       Journal code: 100883449. ISSN: 1531-605X.
PUB. COUNTRY: United States
DOCUMENT TYPE: Journal; Article; (JOURNAL ARTICLE)
LANGUAGE: English
FILE SEGMENT: Priority Journals
ENTRY MONTH: 200002
ENTRY DATE: Entered STN: 20000209
                       Last Updated on STN: 20000209
                       Entered Medline: 20000201

AB   Statistical natural **language** processors have been the focus of
      much research during the past decade. The main advantage of such an
      approach over grammatical rule-based approaches is its scalability to new
      domains. We present a statistical NLP for the domain of radiology and
      report on methods of knowledge acquisition, **parsing**, semantic
      interpretation, and evaluation. Preliminary performance data are given. A
      discussion of the perceived benefit, limitations and future work is
      presented.

ACCESSION NUMBER: 2000032954   MEDLINE
DOCUMENT NUMBER: 20032954   PubMed ID: 10566444
TITLE: Extracting noun phrases for all of MEDLINE.
AUTHOR: Bennett N A; He Q; Powell K; Schatz B R
CORPORATE SOURCE: CANIS-Community Architectures for Network Information
                       Systems, Graduate School of Library and Information
                       Science, University of Illinois at Urbana-Champaign 61820,
                       USA.. nabennet@canis.uiuc.edu
SOURCE: PROCEEDINGS / AMIA SYMPOSIUM, (1999) 671-5.
                       Journal code: 100883449. ISSN: 1531-605X.
PUB. COUNTRY: United States
DOCUMENT TYPE: Journal; Article; (JOURNAL ARTICLE)
LANGUAGE: English
FILE SEGMENT: Priority Journals
ENTRY MONTH: 200002
ENTRY DATE: Entered STN: 20000209
                       Last Updated on STN: 20000209
                       Entered Medline: 20000201

AB   A natural **language parser** that could extract noun
      phrases for all **medical** texts would be of great utility in
      analyzing content for information retrieval. We discuss the extraction of
      noun phrases from MEDLINE, using a general **parser** not tuned
      specifically for any **medical** domain. The noun phrase extractor
      is made up of three modules: tokenization; part-of-speech tagging; noun
      phrase identification. Using our program, we extracted noun phrases from
      the entire MEDLINE collection, encompassing 9.3 million abstracts. Over
      270 million noun phrases were generated, of which 45 million were unique.
      The quality of these phrases was evaluated by examining all phrases from a
      sample collection of abstracts. The precision and recall of the phrases
      from our general **parser** compared favorably with those from three
      other **parsers** we had previously evaluated. We are continuing to
      improve our **parser** and evaluate our claim that a generic
      parser can effectively extract all the different phrases across
      the entire **medical** literature.

ACCESSION NUMBER: 1999122618   MEDLINE
DOCUMENT NUMBER: 99122618   PubMed ID: 9925230
TITLE: Representing information in patient reports using natural
                     language processing and the extensible markup language.
AUTHOR: Friedman C; Hripcsak G; Shagina L; Liu H
CORPORATE SOURCE: Department of Medical Informatics, Columbia University, New
                     York, New York 10032, USA.. friedma@flux.cpmc.columbia.edu
CONTRACT NUMBER: LM05627 (NLM)
     LM06274 (NLM)
SOURCE: JOURNAL OF THE AMERICAN MEDICAL INFORMATICS ASSOCIATION,
                     (1999 Jan-Feb) 6 (1) 76-87.
                     Journal code: 9430800. ISSN: 1067-5027.
PUB. COUNTRY: United States
DOCUMENT TYPE: Journal; Article; (JOURNAL ARTICLE)
LANGUAGE: English
FILE SEGMENT: Priority Journals
ENTRY MONTH: 199902
ENTRY DATE: Entered STN: 19990301
                     Last Updated on STN: 19990301
                     Entered Medline: 19990216
AB   OBJECTIVE: To design a document model that provides reliable and efficient
     access to clinical information in patient reports for a broad range of
     clinical applications, and to implement an automated method using natural
     language processing that maps textual reports to a form consistent
     with the model. METHODS: A document model that encodes structured clinical
     information in patient reports while retaining the original contents was
     designed using the extensible markup language (XML), and a
     document type definition (DTD) was created. An existing natural
     language processor (NLP) was modified to generate output
     consistent with the model. Two hundred reports were processed using the
     modified NLP system, and the XML output that was generated was validated
     using an XML validating parser. RESULTS: The modified NLP system
     successfully processed all 200 reports. The output of one report was
     invalid, and 199 reports were valid XML forms consistent with the DTD.
     CONCLUSIONS: Natural language processing can be used to
     automatically create an enriched document that contains a structured
     component whose elements are linked to portions of the original textual
     report. This integrated document model provides a representation where
     documents containing specific information can be accurately and
     efficiently retrieved by querying the structured components. If manual
     review of the documents is desired, the salient information in the
     original reports can also be identified and highlighted. Using an XML
     model of tagging provides an additional benefit in that software tools
     that manipulate XML documents are readily available.

ACCESSION NUMBER: 1998136430   MEDLINE
DOCUMENT NUMBER: 98136430   PubMed ID: 9475953
TITLE: Dependency parsing for medical
                     language and concept representation.
AUTHOR: Steimann F
CORPORATE SOURCE: Institut fur Rechnergestutzte Wissensverarbeitung,
                     Universitat Hannover, Germany.. steimann@acm.org
SOURCE: ARTIFICIAL INTELLIGENCE IN MEDICINE, (1998 Jan) 12 (1)
                     77-86.
                     Journal code: 8915031. ISSN: 0933-3657.
PUB. COUNTRY: Netherlands
DOCUMENT TYPE: Journal; Article; (JOURNAL ARTICLE)
LANGUAGE: English
FILE SEGMENT: Priority Journals
ENTRY MONTH: 199804
ENTRY DATE: Entered STN: 19980430
                     Last Updated on STN: 19980430

AB   The theory of conceptual structures serves as a common basis for natural
     language processing and **medical** concept representation.
     We present a PROLOG-based formalization of dependency grammar that can
     accommodate conceptual structures in its dependency rules. First results
     indicate that this formalization provides an operational basis for the
     implementation of **medical language parsers**
     and for the design of **medical** concept representation languages.

L4   ANSWER 46 OF 90   MEDLINE
ACCESSION NUMBER: 1998020609   MEDLINE
DOCUMENT NUMBER: 98020609   PubMed ID: 9357738
TITLE: A natural **language parsing** system for
                     encoding admitting diagnoses.
AUTHOR: Haug P J; Christensen L; Gundersen M; Clemons B; Koehler S;
                     Bauer K
CORPORATE SOURCE: LDS Hospital, USA.
CONTRACT NUMBER: HL53427 (NHLBI)
     LM05323 (NLM)
SOURCE: PROCEEDINGS / AMIA ANNUAL FALL SYMPOSIUM, (1997) 814-8.
                     Journal code: 9617342. ISSN: 1091-8280.
PUB. COUNTRY: United States
DOCUMENT TYPE: Journal; Article; (JOURNAL ARTICLE)
LANGUAGE: English
FILE SEGMENT: Priority Journals
ENTRY MONTH: 199712
ENTRY DATE: Entered STN: 19980109
                     Last Updated on STN: 19980109
                     Entered Medline: 19971217
AB   Free-text or natural **language** documents make up an increasing
     part of the computerized **medical** record. While they do provide
     accessible clinical information to health care personnel, they fail to
     support processes that require clinical data coded according to a shared
     lexicon and data structure. We have developed a natural **language**
     parser that converts free-text admitting diagnoses into a coded
     form. This application has proven acceptably accurate in the experimental
     laboratory to warrant a test in the target clinical environment. Here we
     describe an approach to moving this research application into a production
     environment where it can contribute to the efforts of the Health
     Information Services Department. This transition is essential if the
     products of natural **language** understanding research are to
     contribute to patient care in a routine and sustainable way.

L4   ANSWER 50 OF 90   MEDLINE
ACCESSION NUMBER: 97103275   MEDLINE
DOCUMENT NUMBER: 97103275   PubMed ID: 8947647
TITLE: Recognizing noun phrases in **medical** discharge
                     summaries: an evaluation of two natural **language**
                     parsers.
AUTHOR: Spackman K A; Hersh W R
CORPORATE SOURCE: Biomedical Information Communication Center Oregon Health
                     Sciences University, Portland, USA.
CONTRACT NUMBER: U01-LM05879 (NLM)
SOURCE: PROCEEDINGS / AMIA ANNUAL FALL SYMPOSIUM, (1996) 155-8.
                     Journal code: 9617342. ISSN: 1091-8280.
PUB. COUNTRY: United States
DOCUMENT TYPE: Journal; Article; (JOURNAL ARTICLE)
LANGUAGE: English
FILE SEGMENT: Priority Journals
ENTRY MONTH: 199701
ENTRY DATE: Entered STN: 19970219
                     Last Updated on STN: 19990129
                     Entered Medline: 19970128
AB   We evaluated the ability of two natural **language parsers**

, CLARIT and the Xerox Tagger, to identify simple, noun phrases in
medical discharge summaries. In twenty randomly selected discharge
summaries, there were 1909 unique simple noun phrases. CLARIT and the
Xerox Tagger exactly identified 77.0% and 68.7% of the phrases,
respectively, and partially identified 85.7% and 80.8% of the phrases.
Neither system had been specially modified or tuned to the **medical**
domain. These results suggest that it is possible to apply existing
natural **language** processing (NLP) techniques to large bodies of
medical text, in order to empirically identify the terminology
used in **medicine**. Virtually all the noun phrases could be
regarded as having special **medical** connotation and would be
candidates for entry into a controlled **medical** vocabulary.

L4   ANSWER 52 OF 90   BIOSIS   COPYRIGHT 2003 BIOLOGICAL ABSTRACTS INC.
ACCESSION NUMBER: 1995:556620   BIOSIS
DOCUMENT NUMBER: PREV199698570920
TITLE: New trends in natural language processing: Statistical
                      natural language processing.
AUTHOR(S):   Marcus, Mitchell
CORPORATE SOURCE: Dep. Computer Information Sci., Univ. Pennsylvania,
                      Philadelphia, PA 19104-6389 USA
SOURCE: Proceedings of the National Academy of Sciences of the
                      United States of America, (1995) Vol. 92, No. 22, pp.
                      10052-10059.
                      ISSN: 0027-8424.
DOCUMENT TYPE: Article
LANGUAGE: English
AB   The field of natural **language** processing (NLP) has seen a
      dramatic shift in both research direction and methodology in the past
      several years. In the past, most work in computational linguistics tended
      to focus on purely symbolic methods. Recently, more and more work is
      shifting toward hybrid methods that combine new empirical corpus-based
      methods, including the use of probabilistic and information-theoretic
      techniques, with traditional symbolic methods. This work is made possible
      by the recent availability of linguistic databases that add rich
      linguistic annotation to corpora of natural **language** text.
      Already, these methods have led to a dramatic improvement in the
      performance of a variety of NLP systems with similar improvement likely in
      the coming years. This paper focuses on these trends, surveying in
      particular three areas of recent progress: part-of-speech tagging,
      stochastic **parsing**, and lexical semantics.

L4   ANSWER 62 OF 90   MEDLINE
ACCESSION NUMBER: 95037248   MEDLINE
DOCUMENT NUMBER: 95037248   PubMed ID: 7949928
TITLE: A natural language understanding system combining syntactic
                      and semantic techniques.
AUTHOR: Haug P; Koehler S; Lau L M; Wang P; Rocha R; Huff S
CORPORATE SOURCE: Department of Medical Informatics, LDS Hospital, Primary
                      Children's Medical Center, Salt Lake City, Utah.
CONTRACT NUMBER: 5 R01 LM05323 (NLM)
SOURCE: PROCEEDINGS / THE ... ANNUAL SYMPOSIUM ON COMPUTER
                      APPLICATIONS IN MEDICAL CARE, (1994) 247-51.
                      Journal code: 8113685. ISSN: 0195-4210.
PUB. COUNTRY: United States
DOCUMENT TYPE: Journal; Article; (JOURNAL ARTICLE)
LANGUAGE: English
FILE SEGMENT: Priority Journals
ENTRY MONTH: 199412
ENTRY DATE: Entered STN: 19950110
                      Last Updated on STN: 19950110
                      Entered Medline: 19941223
AB   A large proportion of the **medical** record currently available in
      computerized **medical** information systems is in the form of free

text reports. While the accessibility of this source of data is improved through inclusion in the computerized record, it remains unavailable for automated decision support, **medical** research, and management of **medical** delivery systems. Natural **language** understanding systems (NLUS) designed to encode free text reports represent one approach to making this information available for these uses. Below we describe an experimental NLUS designed to **parse** the reports of chest radiographs and store the clinical data extracted in a **medical** data base.

AB   OBJECTIVE: Development of a general natural-**language** processor that identifies clinical information in narrative reports and maps that information into a structured representation containing clinical terms. DESIGN: The natural-**language** processor provides three phases of processing, all of which are driven by different knowledge sources. The first phase performs the **parsing**. It identifies the structure of the text through use of a grammar that defines semantic patterns and a target form. The second phase, regularization, standardizes the terms in the initial target structure via a compositional mapping of multi-word phrases. The third phase, encoding, maps the terms to a controlled vocabulary. Radiology is the test domain for the processor and the target structure is a formal model for representing clinical information in that domain. MEASUREMENTS: The impression sections of 230 radiology reports were encoded by the processor. Results of an automated query of the resultant database for the occurrences of four diseases were compared with the analysis of a panel of three physicians to determine recall and precision. RESULTS: Without training specific to the four diseases, recall and precision of the system (combined effect of the processor and query generator) were 70% and 87%. Training of the query component increased recall to 85% without changing precision.

ENTRY MONTH:	199404
ENTRY DATE:	Entered STN: 19940428
	Last Updated on STN: 19940428
	Entered Medline: 19940421

AB   Information retrieval in large information databases is a
non-deterministic process which needs a sequence of search steps
generally. One of the main problems to which the end-users are faced is to
parse efficiently their questions into the query **language**
that the computer systems allow. Conceptual graphs were initially designed
for natural **language** analysis and understanding. Due to their
closeness to semantic networks, their expressiveness is powerful enough to
be applied to knowledge representation and use by computer systems. This
work demonstrates that conceptual graphs are a suitable means to model the
end-users querieson the basis of the thesaurus and the semantic network of
the UMLS project.

L4   ANSWER 68 OF 90   MEDLINE
ACCESSION NUMBER:	94176892 MEDLINE
DOCUMENT NUMBER:	94176892 PubMed ID: 8130544
TITLE:	Generating MEDLINE search strategies using a librarian
	knowledge-based system.
AUTHOR:	Peng P; Aguirre A; Johnson S B; Cimino J J
CORPORATE SOURCE:	Center for Medical Informatics, Columbia-Presbyterian
	Medical Center, New York 10032.
SOURCE:	PROCEEDINGS / THE ... ANNUAL SYMPOSIUM ON COMPUTER
	APPLICATIONS IN MEDICAL CARE, (1993) 596-600.
	Journal code: 8113685. ISSN: 0195-4210.
PUB. COUNTRY:	United States
DOCUMENT TYPE:	Journal; Article; (JOURNAL ARTICLE)
LANGUAGE:	English
FILE SEGMENT:	Priority Journals
ENTRY MONTH:	199404
ENTRY DATE:	Entered STN: 19940428
	Last Updated on STN: 19940428
	Entered Medline: 19940421

AB   We describe a librarian knowledge-based system that generates a search
strategy from a query representation based on a user's information need.
Together with the natural **language parser** AQUA, the
system functions as a human/computer interface, which translates a user
query from free text into a BRS Onsite search formulation, for searching
the MEDLINE bibliographic database. In the system, conceptual graphs are
used to represent the user's information need. The UMLS Metathesaurus and
Semantic Net are used as the key knowledge sources in building the
knowledge base.

L4   ANSWER 69 OF 90   MEDLINE
ACCESSION NUMBER:	94176829 MEDLINE
DOCUMENT NUMBER:	94176829 PubMed ID: 8130481
TITLE:	Interpreting natural language queries using the UMLS.
AUTHOR:	Johnson S B; Aguirre A; Peng P; Cimino J
CORPORATE SOURCE:	Center for Medical Informatics, Columbia Presbyterian
	Medical Center, New York, NY 10032.
SOURCE:	PROCEEDINGS / THE ... ANNUAL SYMPOSIUM ON COMPUTER
	APPLICATIONS IN MEDICAL CARE, (1993) 294-8.
	Journal code: 8113685. ISSN: 0195-4210.
PUB. COUNTRY:	United States
DOCUMENT TYPE:	Journal; Article; (JOURNAL ARTICLE)
LANGUAGE:	English
FILE SEGMENT:	Priority Journals
ENTRY MONTH:	199404
ENTRY DATE:	Entered STN: 19940428
	Last Updated on STN: 19940428
	Entered Medline: 19940421

AB   This paper describes AQUA (A QUery Analyzer), the natural **language**

front end of a prototype information retrieval system. AQUA translates a user's natural **language** query into a representation in the Conceptual Graph formalism. The graph is then used by subsequent components to search various resources such as databases of the **medical** literature. The focus of the **parsing** method is on semantics rather than syntax, with semantic restrictions being provided by the UMLS Semantic Net. The intent of the approach is to provide a method that can be emulated easily in applications that require simple natural **language** interfaces.

L4   ANSWER 71 OF 90   MEDLINE
ACCESSION NUMBER: 93230105   MEDLINE
DOCUMENT NUMBER: 93230105   PubMed ID: 8472004
TITLE: UMLS knowledge for biomedical language processing.
AUTHOR: McCray A T; Aronson A R; Browne A C; Rindflesch T C; Razi A; Srinivasan S
CORPORATE SOURCE: Computer Science Branch, National Library of Medicine, Bethesda, MD 20894.
SOURCE: BULLETIN OF THE MEDICAL LIBRARY ASSOCIATION, (1993 Apr) 81 (2) 184-94.
                        Journal code: 0421037. ISSN: 0025-7338.
PUB. COUNTRY: United States
DOCUMENT TYPE: Journal; Article; (JOURNAL ARTICLE)
LANGUAGE: English
FILE SEGMENT: Priority Journals
ENTRY MONTH: 199305
ENTRY DATE: Entered STN: 19930604
                        Last Updated on STN: 20021030
                        Entered Medline: 19930520
AB   This paper describes efforts to provide access to the free text in biomedical databases. The focus of the effort is the development of SPECIALIST, an experimental natural **language** processing system for the biomedical domain. The system includes a broad coverage **parser** supported by a large lexicon, modules that provide access to the extensive Unified **Medical Language** System (UMLS) Knowledge Sources, and a retrieval module that permits experiments in information retrieval. The UMLS Metathesaurus and Semantic Network provide a rich source of biomedical concepts and their interrelationships. Investigations have been conducted to determine the type of information required to effect a map between the **language** of queries and the **language** of relevant documents. Mappings are never straightforward and often involve multiple inferences.

L4   ANSWER 72 OF 90   MEDLINE   DUPLICATE 13
ACCESSION NUMBER: 93225868   MEDLINE
DOCUMENT NUMBER: 93225868   PubMed ID: 8469163
TITLE: Semantic analysis of **medical** records.
AUTHOR: Rasmussen J E; Bassoe C F
CORPORATE SOURCE: Promed Institute, Bergen, Norway.
SOURCE: METHODS OF INFORMATION IN MEDICINE, (1993 Feb) 32 (1) 66-72.
                        Journal code: 0210453. ISSN: 0026-1270.
PUB. COUNTRY: GERMANY: Germany, Federal Republic of
DOCUMENT TYPE: Journal; Article; (JOURNAL ARTICLE)
LANGUAGE: English
FILE SEGMENT: Priority Journals
ENTRY MONTH: 199305
ENTRY DATE: Entered STN: 19930521
                        Last Updated on STN: 19930521
                        Entered Medline: 19930507
AB   A program (LogStory) is described that was developed for the automatic semantic analysis of clinical narratives, stored in a computerized problem-oriented **medical** record (PROMED). The diagnoses were written in a free-text format during consultation, and later collected

into diagnostic classes, e.g., diseases. A lexical **parser**
automatically created dictionaries from the clinical narrative associated
with each disease. Automatic (fuzzy) set operations were performed on the
words associated with each class. The manifestations of 16 diseases were
automatically extracted by pairwise operations on the word sets. The
correlation between diseases and corresponding signs, symptoms and
treatment was highly significant (p < 0.001). Applying the difference
operation on diseases with disjunct sets of clinical findings allowed the
recovery of disease-specific knowledge. The evolution of a disease was
accounted for, and the system was able to generalize its findings. The
PROMED-LogStory concept enables the processing of natural language
and may be a powerful tool for knowledge acquisition and clinical
research.

AB   For **medical** records, the challenge for the present decade is
      Natural **Language** Processing (NLP) of texts, and the construction
      of an adequate Knowledge Representation. This article describes the
      components of an NLP system, which is currently being developed in the
      Geneva Hospital, and within the European Community's AIM programme. They
      are: a Natural **Language** Analyser, a Conceptual Graphs Builder, a
      Data Base Storage component, a Query Processor, a Natural **Language**
      Generator and, in addition, a Translator, a Diagnosis Encoding System and
      a Literature Indexing System. Taking advantage of a closed domain of
      knowledge, defined around a **medical** specialty, a method called
      proximity processing has been developed. In this situation no
      parser of the initial text is needed, and the system is based on
      semantical information of near words in sentences. The benefits are: easy
      implementation, portability between languages, robustness towards
      badly-formed sentences, and a sound representation using conceptual
      graphs.

AB   This paper describes MetaIndex, an automatic indexing program that creates
     symbolic representations of documents for the purpose of document
     retrieval. MetaIndex uses a simple transition network **parser** to
     recognize a **language** that is derived from the set of main
     concepts in the Unified **Medical Language** System
     Metathesaurus (Meta-1). MetaIndex uses a hierarchy of **medical**
     concepts, also derived from Meta-1, to represent the content of documents.
     The goal of this approach is to improve document retrieval performance by
     better representation of documents. An evaluation method is described, and
     the performance of MetaIndex on the task of indexing the Slice of Life
     medical image collection is reported.

L4   ANSWER 78 OF 90   MEDLINE   DUPLICATE 16
AB   Much information about patients is stored in free text. Hence, the
     computerized processing of **medical language** data has
     been a well-known goal of **medical** informatics resulting in
     different paradigms. In Gottingen, a **Medical** Text Analysis
     System for German (abbr. MediTAS) has been under development for some
     time, trying to combine and to extend these paradigms. This article
     concentrates on the automated syntax analysis of German **medical**
     utterances. The investigated text material consists of 8,790 distinct
     utterances extracted from the summary sections of about 18,400
     cytopathological findings reports. The **parsing** is based upon a
     new approach called Left-Associative Grammar (LAG) developed by Hausser.
     By extending considerably the LAG approach, most of the grammatical
     constructions occurring in the text material could be covered.

L4   ANSWER 79 OF 90   MEDLINE

AB   Over the past several years our research efforts have been directed toward
      the identification of natural **language** processing methods and
      techniques for improving access to biomedical information stored in
      computerized form. To provide a testing ground for some of these ideas we
      have undertaken the development of SPECIALIST, a prototype system for
      parsing and accessing biomedical text. The system includes
      linguistic and biomedical knowledge. Linguistic knowledge involves rules
      and facts about the grammar of the **language**. Biomedical
      knowledge involves rules and facts about the domain of biomedicine. The
      UMLS knowledge sources, Meta-1 and the Semantic Network, as well as the
      UMLS test collection, have recently contributed to the development of the
      SPECIALIST system.

L4   ANSWER 80 OF 90   EMBASE   COPYRIGHT 2003 ELSEVIER SCI. B.V.
AB   This paper shows how regularities of **language** usage within a
      narrow subject area (sublanguage) are used in computerized informational
      analysis of free-text input. Documents are processed by the NYU Linguistic
      String Project (LSP) **parsing** system, which uses a computer
      grammar of English, a detailedly coded lexicon, English transformations to
      regularize syntactic structures, and a semantic component based on
      sublanguage co-occurrence patterns. The workings of the system and an
      application to free-text **medical** douments are described. Recent
      work on French **medical** documents is included.

L4   ANSWER 83 OF 90   MEDLINE
AB   The organization of brain processes leading to **language** and
      movement show important parallels with one another and also express
      important aspects of **biological** organization in general. Four
      major differences between **biological** processes and their
      commonly proposed analogues, machine processes, are as follows. 1)

Reduction is not simplification in **biological** analysis; rather
the subsystems that result from separation of parts of a
biological system are themselves complex, often potentially
viable, systems. 2) Machine processes are typically generalized, or, if
specialized, they are specialized by connecting general-type subsystems in
special ways. But **biological** systems are typically specialized
at many levels, both in subsystems and their connections. 3) The history
of a **biological** system is often an intimate and inseparable part
of its structure. Furthermore **biological** systems never develop
alone or de novo. Not only do they develop in clusters of contemporaries,
they also develop in the presence of an older generation and a "culture."
4) Not only do formal logics have some constraints that **biological**
minds may not have (e.g., internal consistency and universality), formal
logics require descriptions of qualitative phenomena in a **language**
that is inadequate and (as a deeper issue) may always require
parsing a meaningful whole into approximate parts (e.g., as in
writing this abstract). Instances of contrasts between **biological**
systems and machine-type systems are seen in **language** and
movement phenomena, such as embodying a distinction between purposes and
causes and having flexibly reorganizable subassemblies, multiple goals,
and motor equivalence.

L4   ANSWER 84 OF 90   EMBASE   COPYRIGHT 2003 ELSEVIER SCI. B.V.
ACCESSION NUMBER: 84151578   EMBASE
DOCUMENT NUMBER: 1984151578
TITLE: **Biological** processing.
AUTHOR: Bellman K.L.; Walter D.O.
CORPORATE SOURCE: Crump Institute for Medical Engineering, University of
                       California, Los Angeles, CA 90024, United States
SOURCE: American Journal of Physiology - Regulatory Integrative and
                       Comparative Physiology, (1984) 15/6 (R860-R867).
                       CODEN: AJPRDO
COUNTRY: United States
DOCUMENT TYPE: Journal
FILE SEGMENT: 002   Physiology
                       027   Biophysics, Bioengineering and Medical
                                Instrumentation
                       008   Neurology and Neurosurgery
LANGUAGE: English
AB   The organization of brain processes leading to **language** and
       movement show important parallels with one another and also express
       important aspects of **biological** organization in general. Four
       major differences between **biological** processes and their
       commonly proposed analogues, machine processes, are as follows. 1)
       Reduction is not simplification in **biological** analysis; rather
       the subsystems that result from separation of parts of a
       biological system are themselves complex, often potentially
       viable, systems. 2) Machine processes are typically generalized, or, if
       specialized, they are specialized by connecting general-type subsystems in
       special ways. But **biological** systems are typically specialized
       at many levels, both in subsystems and their connections. 3) The history
       of a **biological** system is often an intimate and inseparable part
       of its structure. Furthermore **biological** systems never develop
       alone or de novo. Not only do they develop in clusters of contemporaries,
       they also develop in the presence of an older generation and a 'culture'.
       4) Not only do formal logics have some constraints that **biological**
       minds may not have (e.g., internal consistency and universality), formal
       logics require descriptions of qualitative phenomena in a **language**
       that is inadequate and (as a deeper issue) may always require
       parsing a meaningful whole into approximate parts (e.g., as in
       writing this abstract). Instances of contrasts between **biological**
       systems and machine-type systems are seen in **language** and
       movement phenomena, such as embodying a distinction between purposes and
       causes and having flexibly reorganizable subassemblies, multiple goals,

and motor equivalence.

ACCESSION NUMBER: 80193743   EMBASE
DOCUMENT NUMBER: 1980193743
TITLE: Profile of a dictionary compiled from scanning over one
                     million words of surgical pathology narrative text.
AUTHOR: Wong R.L.; Reno J.D.; Hain T.C.; et al.
CORPORATE SOURCE: Dept. Pathol., Abraham Lincoln Sch. Med., Univ. Illinois,
                     Chicago, Ill. 60612, United States
SOURCE: Computers and Biomedical Research, (1980) 13/4 (382-398).
                     CODEN: CBMRB7
COUNTRY: United States
DOCUMENT TYPE: Journal
FILE SEGMENT: 027   Biophysics, Bioengineering and Medical
                              Instrumentation
LANGUAGE: English

AB   An anatomic pathology natural **language** dictionary (LEXICON) has
     evolved over a 9-yr period, a result of scanning over one million words of
     narrative text from tissue examination request forms and surgical
     pathology reports. The text is **parsed** into individual words
     which are looked up in LEXICON and flagged by action codes which determine
     usage in constructing a KWIC index file and an on-line database
     retrievable by keywords. The LEXICON now resides on an IBM 370/168 system
     and has survived several transfers between computer systems. An update
     program is used after each batch of narrative text is scanned to modify
     LEXICON. LEXICON now contains 24,228 **medical** and nonmedical
     terms, 24.8% are errors (misspellings), 45.9% are keywords retrievable on
     and off line. 52.2% of the words are cross-referenced to a supplementary
     word. A preliminary study shows that many of the 'nonmedical' terms in
     LEXICON carry significant **medical** information, and that there is
     considerable overlap of **medical** words among LEXICON, SNOMED, and
     ICDA-8. The authors' LEXICON appears to be an intermediate step in the
     process of evolving an algorithm capable of 'understanding'
     medical narrative text.